712CD

75TH MORSS CD Cover Page

75TH MORSS
US Naval Academy
Annapolis, Maryland
June 12-13-14 2007
Through Multiple Disciplines, Analytical Power

If you would like your presentation included in the 75th MORSS Final Report CD it must :

1. *Be unclassified, approved for public release, distribution unlimited, and is exempt from U.S. export licensing and other export approvals including the International Traffic in Arms Regulations (22CFR120 et seq.);*
2. Include MORS Form 712CD as the first page of the presentation;
3. Have an approved MORS form 712 A/B and
4. Be turned into the MORS office no later than: **DEADLINE: 14 June 2007 (Late submissions will not be included.)**

Author Request (To be completed by applicant) - The following author(s) request authority to disclose the following presentation in the MORSS Final Report, for inclusion on the MORSS CD and/or posting on the MORS web site.

Name of Principal Author and all other author(s): Richard Shaffer

Principal Author's Organization and address:
G1-CPTD WAFO
2461 Eiisenhower Avenuew
Hoffman I, Room 172
Alexandria, VA 22331

Phone: 703-325-1585

Fax: 703-325-3894

Email: Richard.a.shaffer1@us.army.mil

Please use the same title listed on the 75TH MORSS Disclosure Form 712 A/B. If the title of the presentation has changed please list both.)

Original title on 712 A/B:
Diagnosing Long Running LP Models in the Civilian Forecasting System

If the title was revised please list the original title above and the revised title here:

PRESENTED IN:

WORKING GROUP:	20	**DEMONSTRATION:**	
COMPOSITE GROUP:		**POSTER:**	
SPECIAL SESSION 1:		**TUTORIAL:**	
SPECIAL SESSION 2:		**OTHER:**	**Information Paper**
SPECIAL SESSION 3:			

This presentation is believed to be: ***Unclassified, approved for public release, distribution unlimited, and is exempt from U.S. export licensing and other export approvals including the International Traffic in Arms Regulations (22CFR120 et seq.)***

UNCLASSIFIED

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 JUN 2007**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Diagnosing Long Running LP Models in the Army Civilian Forecasting System**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **G1-CPTD WAFO 2461 Eiisenhower Avenuew Hoffman I, Room 172 Alexandria, VA 22331**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM202526. Military Operations Research Society Symposium (75th) Held in Annapolis, Maryland on June 12-14, 2007, The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**37**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Diagnosing Long Running LP Models in the Army Civilian Forecasting System

June 2007

UNCLASSIFIED

Agenda

- Purpose
- CIVFORS Overview
- LP Formulation
- Runtime Analysis
- Solution Strategies
- Further Study

Purpose

- The Army Civilian Forecasting System (CIVFORS) was developed in 1987 to help align the civilian workforce with Army structure by Command
- An increase in custom models has led to dramatic differences in runtime for production forecasts
- Improving runtime and managing user expectations requires a greater understanding of runtime drivers

CIVFORS Overview

UNCLASSIFIED

CIVFORS

- A Workforce Planning System
  - Forecasts strength and staffing actions (gains, losses, migrations)
  - Web enabled
  - Flexible design
  - What-if Analysis
  - Goal Setting (Prescriptive Modeling)

History

- Developed by Army in 1987
- Derived from active Army personnel forecasting systems
- Methodology adapted for civilian modeling
- PC version developed in 1998
- Flexible system developed in 2000
- Web enabled in 2003

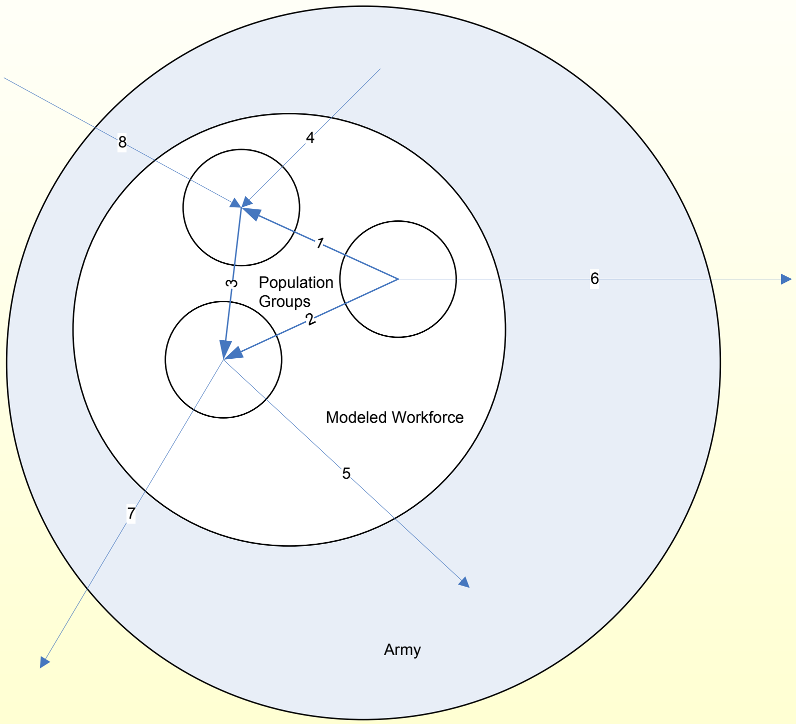
Model Scope

- Army Civilian Corps
  - A career life cycle model
  - Population groups are the state variables
  - Aging and Staffing actions (Gains, losses and migrations) are the model dynamics
- The state of the system is evaluated by the size and distribution (profile) of the population groups
- Dynamics change the state of the system over time
- Deterministic and Linear

Population Groups

- Groups employees with similar characteristics (age, year of service, occupation, etc.)
- Forecasts are computed on a group by group basis
- Groups are assumed to have similar loss and migration rates

Group Dynamics



- Arcs 1-3 represent internal migrations
- Arcs 4-5 represent transfers within Army but outside of the scope of the modeled workforce
- Arcs 6-8 represent separations and new hires respectively

UNCLASSIFIED

Data Sources

- CIVFORS warehouses 5 years of the latest historical data from HQ ACPERS
  - Stores quarterly snapshots of employee records
  - Accumulates Nature of Action records from SF 50s
- Builds special queries for your workforce
  - Tabulates population group statistics
  - Tabulates gains and losses to each group
  - Tabulates aging and migration actions by comparing employee records quarter by quarter

Change Tracking

	Current Quarter				Last Quarter			
Employee ID	Age	Grade	Gender	Occupation	Age	Grade	Gender	Occupation
12345	23	7	F	2210				
12346					58	13	M	0343
12347	43	11	F	0301	43	9	F	0301
12348	48	11	M	1515	48	11	M	1515
12349	32	9	F	0201	31	9	F	0201
12350	39	12	M	0801	39	12	M	0854
…								

*Data does not represent real Army employees and is for demonstration purposes only

- Each area highlighted is tabulated as either a gain, loss, migration or aging action

Historical Rates

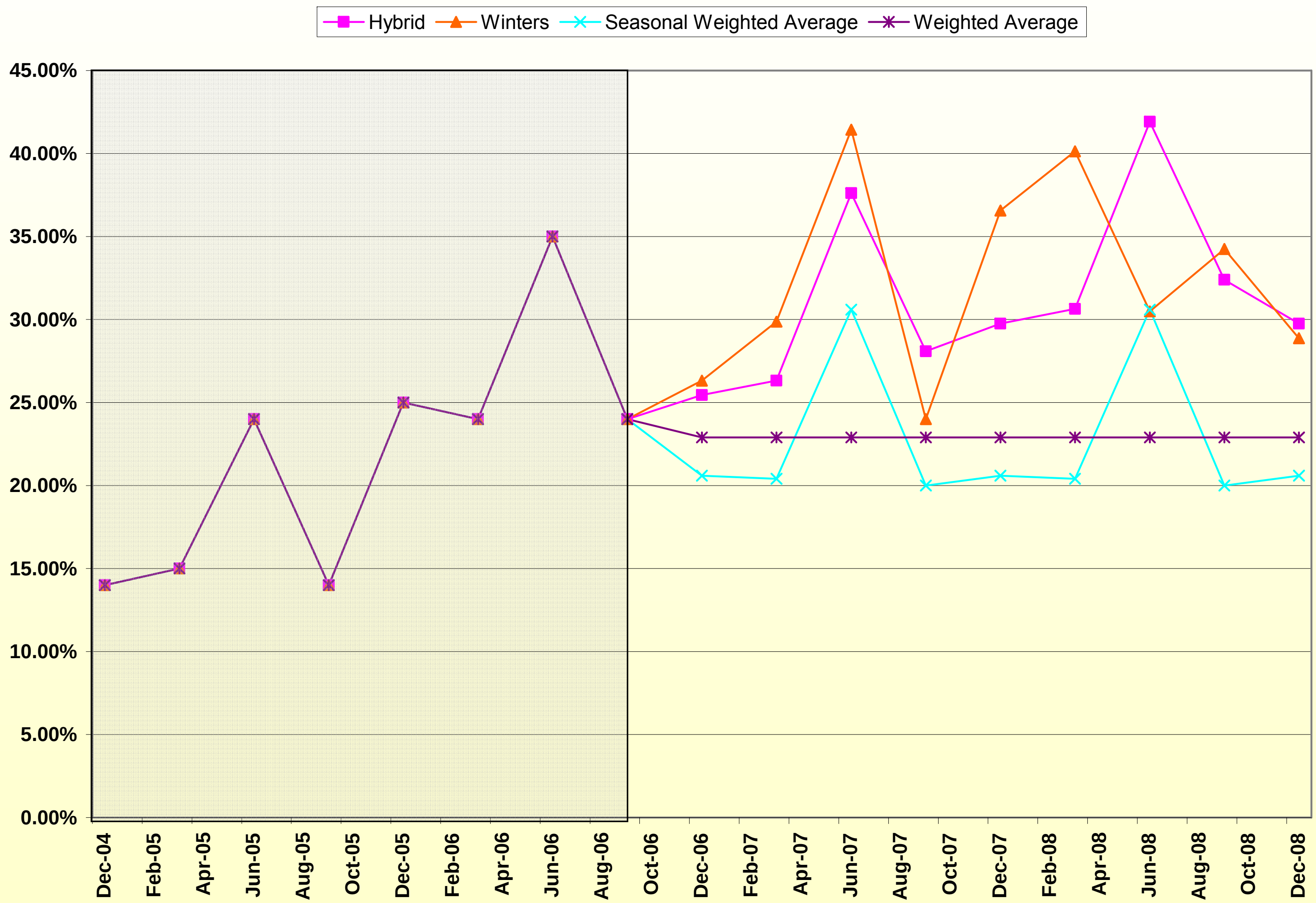
- The life cycle progression is:
  - Age → Hire → Promote → Separate (resign/retire)
- Rates, for each group in a fiscal quarter, are:

Population Group: 40-45 year olds, Grd 7, Females						
	Aging		Hires	Promotions		Separations
Start = 100	Out	In		Out	In	
Tally	15	12	10	5	8	9
Rate Formula	$\frac{15}{100}$			$\frac{5}{100-15+12+10}$		$\frac{9}{100-15+12+10-5+8}$
Rate	15%			4.7%		8.2%

*Data does not represent real Army employees and is for demonstration purposes only

- Incoming group migrations are computed by percent distribution
- Gains (hires) are forecasted as counts per quarter

Forecasting Rates



UNCLASSIFIED

Data Smoothing

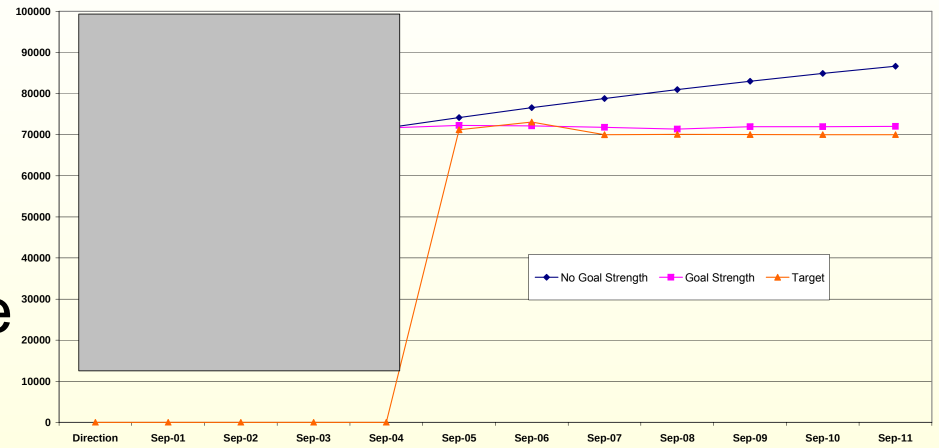
- Weighting Schemes
  - Weights historical years (e.g. .40, .25, .15, .10, .10) to emphasize periods thought to be more representative of the future
- Rate Blending
  - Considers higher level group statistics (e.g. rates by gender only) when computing rates for small population groups
- Outlier detection
  - Excludes rate values that are extreme (with regard to the median over history) when computing the forecast

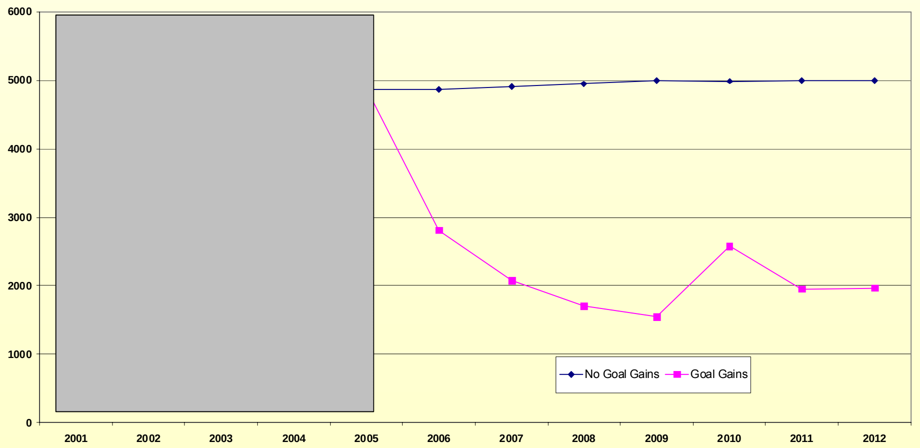
Final Forecasting

- Strength is forecasted with or without manpower goals
- Forecasts without goals are based on applying life cycle rates to a starting inventory of population groups
- Forecasts with goals are developed as the solution of a multi-period, goal linear program

Goal vs NoGoal Forecasts

- Forecasts balance near term and long term manpower goals
- Forecasts are prescriptive and not bound to historical trends
- Gains are more commonly used than other staffing actions like as decision variables





Goal Optimization

CIVFORS Objective

- In Goal Mode, CIVFORS minimizes the deviation between strength and manpower targets
- Manpower targets are derived from
  - The Structure and Manpower Allocation System (SAMAS)
  - The current population profile
- Custom targets created through the rate/target editor user control

Formulation Basics

- The linear program consists of 3 main parts
  - Objective function minimizes strength deviation variables
  - Life Cycle equations compute strength through time
  - Targeting equations compute strength deviation
- Equations are generated using AMPL
- Index sets track
  - Population groups
  - Life cycle actions {gains, migrations, losses}

Variables

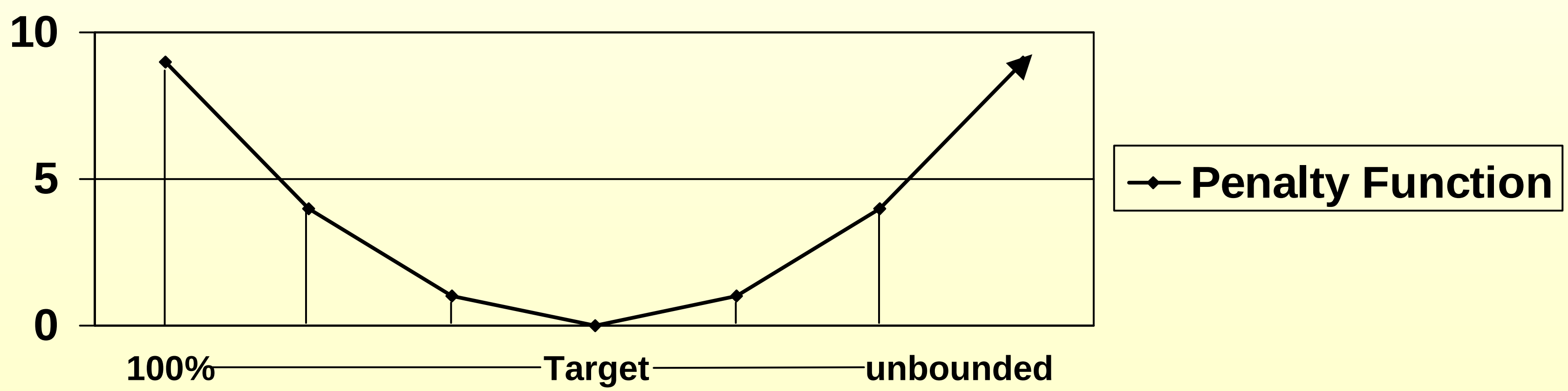
- Strength Accounting
  - Future strength {p in population,mt in modeltime}
  - Future gains {g in optimizedgains, p in population, mt in modeltime}
  - Future losses {l in optimizedlosses, p in population, mt in modeltime}
  - Future migrations {m in optimizedmigrations, p1 in population, p2 in population, mt in modeltime:(m,p1,p2) in migration_factorSet}
  - Life cycle {p in population, 0..numtime+1+nummigration+1,mt in modeltime}
- Strength Targeting
  - TargetShortage{targetpopulation,1..NumberofModelTimePeriods ,BoundPercent}
  - TargetSurplus{targetpopulation,1..NumberofModelTimePeriods, BoundPercent}

Life Cycle Equations

- Life cycle variables track strength at each increment
  - Startup {p in population,mt in modeltime}:
  - TimeDimension_Increment {t in timetrans, p in population, d in dims, mt in modeltime: (t,d) in timetrandims}:
  - Gain_Transactions { p in population, mt in modeltime }:
  - Migration_Transactions { m in migrationtrans, p in population, d in dims, mt in modeltime:(m,d) in migrationtrandims }:
  - Loss_Transactions { p in population, mt in modeltime }:
  - EndTieIn {p in population,mt in modeltime}:
- Strength in a period is equal to the value of the final Life cycle variable

Targeting Equations

- Strength = Target + Shortage – Surplus
- Multiple Shortage and Surplus slacks form a piecewise linear penalty function around each manpower target



UNCLASSIFIED

Runtime Analysis

UNCLASSIFIED

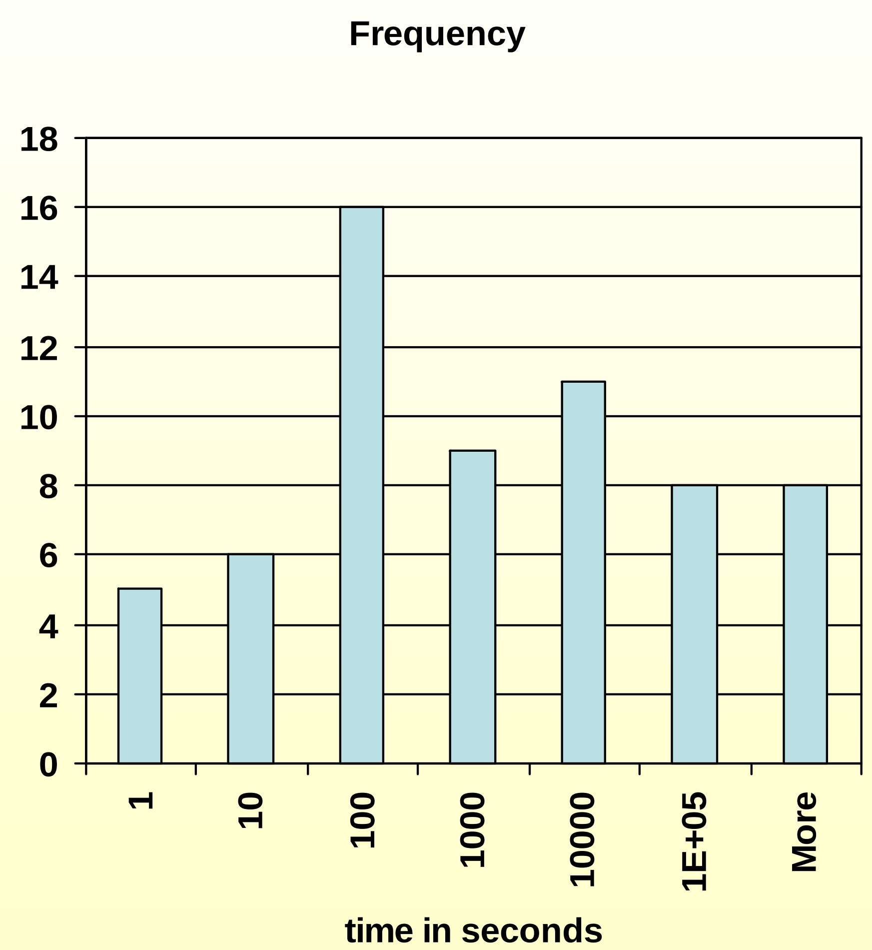
System Specifications

- HP rp3440
- 64-bit HPUX OS
- 2x 1GHz PA-RISC
- 12 GB memory
- All models were formulated and solved using Ilog's AMPL/CPLEX suite (version 10.100 for Unix)
- Ran Primal Simplex without Presolve Option

Baseline Production Runtime

- Average = 22,373.32
- Median = 131.785
- Mode = 172,854*

- Wide variance impacts end user acceptance and use of the system
- *Runtime limited in production to approximately 48 hours



Model Parameters

- CIVFORS limits models based on the number of population groups
- Based on analysis, population is highly correlated with LP size, but less so to actual runtime

	Number of Population Groups	Number of Initial LP Variables	Number of Initial LP Constraints	Number of Initial LP non-zeros	Number of Initial LP Objective Function non-zeros	Solve Time
Number of Population Groups	1					
Number of Initial LP Variables	0.969601	1				
Number of Initial LP Constraints	0.976022	0.996405	1			
Number of Initial LP non-zeros	0.964303	0.980681	0.983575	1		
Number of Initial LP Objective Function non-zeros	0.903344	0.963876	0.938128	0.92516	1	
Solve Time	0.463979	0.523059	0.499795	0.561471	0.573461	1

Initial Regression Results

- LP Matrix Density is a key factor in predicting runtime due to degeneracy and numerical instability
- What contributes to increased density?
- What other factors are important?

Used Natural Log Transformation **R squared = .878**

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-26.318475	4.033025	-6.52574	1E-08	-34.3663	-18.2707	-34.3663	-18.2707
Number of Population Groups	-0.4236863	0.615343	-0.68854	0.493457	-1.65158	0.804211	-1.65158	0.804211
Number of Initial LP Variables	-4.4123977	4.246927	-1.03896	0.302502	-12.887	4.062214	-12.887	4.062214
Number of Initial LP Constraints	-4.7111516	2.933737	-1.60585	0.11294	-10.5653	1.143029	-10.5653	1.143029
Number of Initial LP non-zeros	9.57172491	1.077077	8.886764	5.41E-13	7.422452	11.721	7.422452	11.721
Number of Initial LP Objective Function non-zeros	2.04402933	1.462176	1.397937	0.166676	-0.8737	4.961756	-0.8737	4.961756

Additional Factors Studied

- Data collected for 220 forecasts

Size of Migration Set	Minimum Rate Value
Size of Gain Set	Initial Sparsity Ratio (zero Strength Cells/ Population Groups)
Size of Loss Set	AMPL Presolve Option On (Binary variable)
Size of Migration Factor Set	CPLEX Presolve Option On (Binary variable)
Size of Target Set	Rate Reduction Factor (CIVFORS heuristic)
Total Starting Strength	Number of Reduced Variables
Size of Aggregate Target Set	Number of Reduced Constraints
Total Target Row (Binary variable)	Number of Reduced non-zeros
Number of Optimized Variables	Minimum Migration Factor
Number of zero Strength Cells	Number of zero Targets
RHS Ratio (largest to smallest value)	Maximum Gain Bound
Number of Model Time Periods	Solve Type (primal or dual)

Stepwise Regression

Used Natural Log Transformation R squared = .9162

Variable	Estimate	Error	Type II SS	F Value	Pr > F
Intercept	-23.453	1.49564	383.7786	245.89	<.0001
Size of Migration Set	-0.59833	0.12604	35.17066	22.53	<.0001
Size of Target Set	0.34977	0.06475	45.53614	29.18	<.0001
Size of Loss Set	0.41996	0.12093	18.82345	12.06	0.0007
Size of Migration Factor Set	1.02995	0.16694	59.40812	38.06	<.0001
Size of Aggregate Target Set	0.26169	0.04531	52.07192	33.36	<.0001
Number of Model Time Periods	2.84243	0.53046	44.81464	28.71	<.0001
Solve Type	-1.09972	0.3921	12.27752	7.87	0.0058
Presolve Option On	-0.90403	0.357	10.00877	6.41	0.0124
Rate Reduction Factor	0.64309	0.34923	5.29257	3.39	0.0677
Number of Reduced non-zeros	1.07249	0.13622	96.74319	61.98	<.0001
Initial Sparsity Ratio	0.27006	0.16379	4.24319	2.72	0.1014

Interpretation

- Regression results can be used to bin runtime drivers into three types
  - Factors that can be addressed with Parameter Changes
  - Factors that can be addressed with Model Changes
  - Factors that can be used to manage user expectations

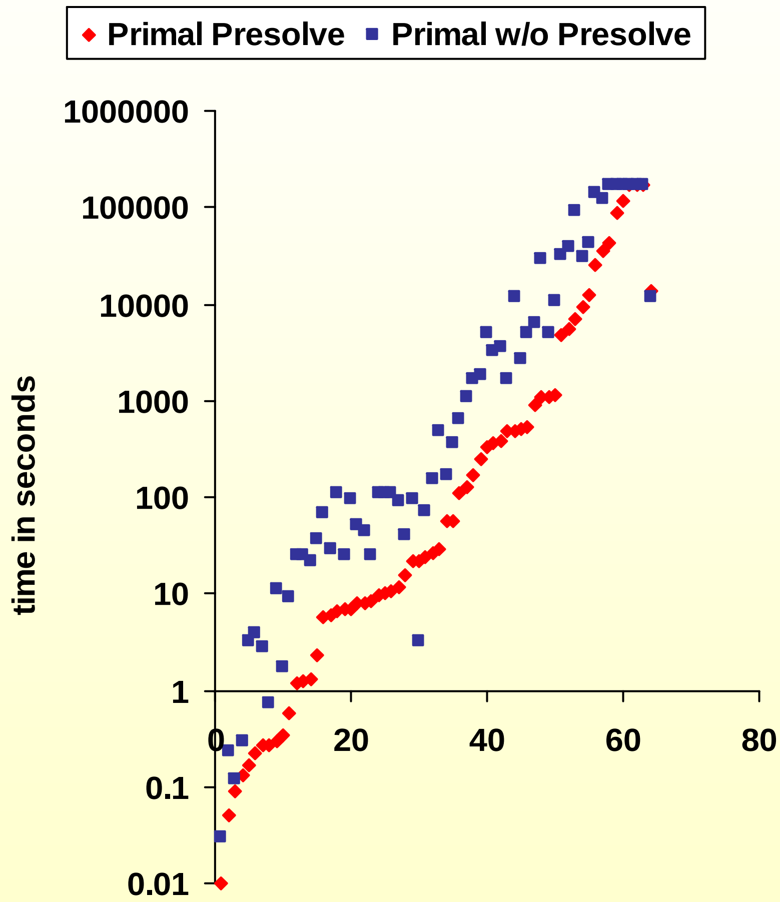
Type 1	Type 2	Type 3
Solve Type	Size of Migration Set	Size of Target Set
Presolve Option On	Size of Migration Factor Set	Size of Aggregate Target Set
	Rate Reduction Factor	Number of Model Time Periods
	Size of Loss Set	Number of Reduced non-zeros
		Initial Sparsity Ratio

Work in Progress

UNCLASSIFIED

CPLEX Presolve

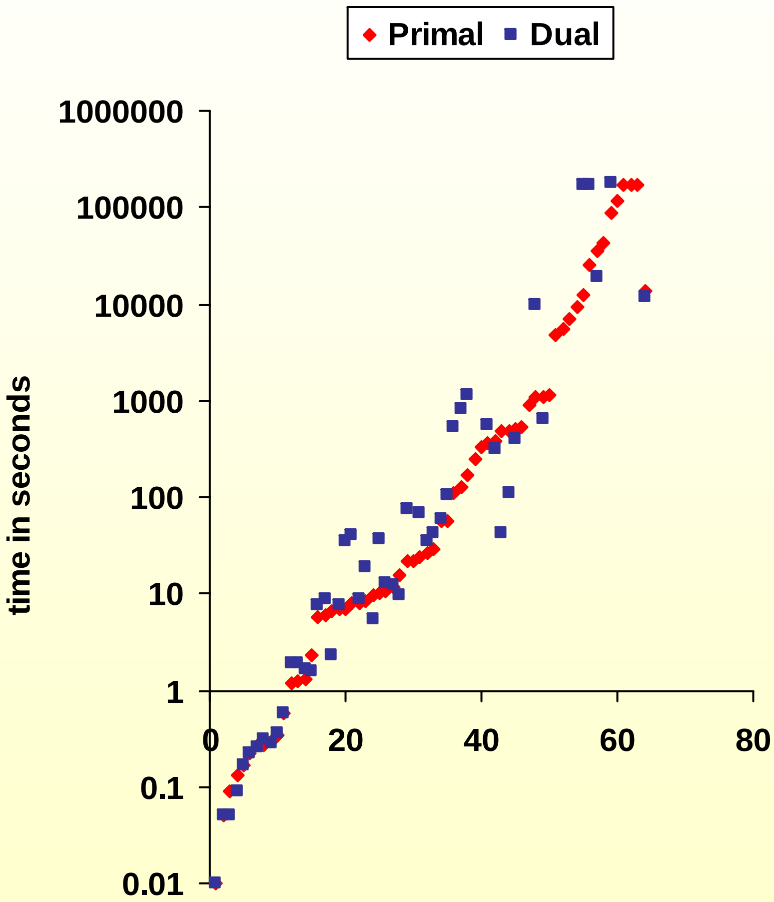
- Using CPLEX Presolve option consistently out performs Primal Simplex on the original formulation by an average of 8 to 1



UNCLASSIFIED

Primal versus Dual

- For comparable runs with Presolve on, Primal out performed Dual Simplex
- However
  - When primal is faster it is 46% faster on average
  - When dual is faster it is 69% faster on average



New Modeling Approaches

- Rate Reduction Methods
  - Utilizing sparsity to limit the number of small rates
  - Adjusting rate blending techniques for migration rates
  - Migration Pooling
    - Creating large, distribution nodes to manage migrations
- Utilizing Barrier methods on low density models

Migration Pooling Results

- Preliminary tests show a more than 10 fold reduction in runtime for comparable models
  - Greater reductions can be achieved when optimizing migration distribution patterns
- Barrier algorithm shows reduced effects from degeneracy and more consistent results across model types
- Barrier Cholesky Factorization statistic is a reliable measure of expected runtime

Further Study

- Extensive testing of migration pooling formulation and Barrier algorithm
- Develop algorithm modifications to support rate reduction methods
- Develop early warning feedback mechanism to user interface
- Develop interface with CPLEX software to manage model run priority